DEPARTMENT OF COMMERCE
SINCLAIR WEEKS, Secretary

WEATHER BUREAU
F. W. REICHELDERFER, Chief

MONTHLY WEATHER REVIEW

JAMES E. CASKEY, JR., Editor

Volume 85
Number 10

OCTOBER 1957

Closed December 15, 1957
Issued January 15, 1958

NUMERICAL METHODS IN WEATHER PREDICTION: I. THE BALANCE EQUATION*

FREDERICK G. SHUMAN, U. S. Weather Bureau

Joint Numerical Weather Prediction Unit, Suitland, Md.

[Manuscript received June 28, 1957; revised October 25, 1957]

ABSTRACT

Two methods of solving the balance equation are outlined. Both methods have been used successfully on a daily operational basis at the Joint Numerical Weather Prediction Unit for a period of more than a year. Solutions were on the operational grid of 30 x 34 points spaced at 381-km. intervals.

1. INTRODUCTION

The Joint Numerical Weather Prediction Unit has been solving the balance equation on a daily operational basis for over a year on a grid covering more than 100,000,000 km.^{2} of the earth's surface (Shuman [1]). With this record of operations, it may safely be concluded that the methods used will never fail to yield a solution nor behave badly for any meteorological data. In view of difficulties anticipated in the literature concerning the solution of the balance equation on grids comparable in size to that used by the Unit it is felt that a record of our methods will be of some interest.

2. THE BALANCE EQUATION IN POLAR STEREOGRAPHIC SPACE

Perhaps the most common form of the balance equation is, in tangent plane coordinates,

$$f(\psi_{xx}+\psi_{yy})-2(\psi_{xy}^2-\psi_{xx}\psi_{yy}) +(\psi_x f_x+\psi_y f_y)-g(z_{xx}+z_{yy})=0 \quad (1)$$

where f is the Coriolis parameter, ψ is the stream function, g is acceleration of gravity, and z is the height of an isobaric surface. The usual transformation of equation (1) onto the polar stereographic projection is

$$m^2 f(\psi_{xx}+\psi_{yy})-2m^4(\psi_{xy}^2-\psi_{xx}\psi_{yy}) +m^2(\psi_x f_x+\psi_y f_y)-m^2 g(z_{xx}+z_{yy})=0 \quad (2)$$

where m is the map scale factor,

$$m=\frac{\rho_e^2+\rho^2}{2\rho_e r},$$

ρ is radial distance on the projection from the pole, ρ_e is radial distance on the projection from pole to equator, and r is the radius of the earth.

Equation (1), being in tangent plane coordinates, is not a rigorous expression of the dynamic laws, nor is equation (2) a rigorous transformation of equation (1) into polar stereographic space. Because serious questions have arisen in the past concerning possible inconsistencies between the common forms of the balance equation and the true physical laws the forms are meant to express, the balance equation has been derived from the equations of motion in spherical coordinates, and transformed rigorously into Cartesian coordinates on the polar stereographic projection. The derivation and transformation being lengthy, they will not be reproduced here for lack

*This is the first of a series of three articles on the subject of numerical methods of prediction used by the Joint Numerical Weather Prediction Unit.

of space. The resulting form is

$$\frac{1}{2}[m^2(\psi_{xx}+\psi_{yy})+f]^2-\frac{m^4}{2}\left[\psi_{xx}-\psi_{yy}+\underline{\frac{4\rho}{\rho_e^2+\rho^2}\cdot\frac{-y\psi_y+x\psi_x}{\rho}}\right]^2-$$
$$\frac{m^4}{2}\left[-2\psi_{xy}+\underline{\frac{4\rho}{\rho_e^2+\rho^2}\cdot\frac{-x\psi_y-y\psi_x}{\rho}}\right]^2+$$
$$m^2\left[\psi_y f_y+\psi_x f_x-\underline{\frac{\psi_x^2+\psi_y^2}{r^2}}\right]-\left[m^2 g(z_{xx}+z_{yy})+\frac{1}{2}f^2\right]=0$$

where x and y are measured on the projection from the pole. The underlined terms, which are not implicit in the common form (2), are small. With the view to assessing them for integrated consistent effects, the balance equation was solved for one case with them included. Their effects proved trivial, so they will be ignored. We have then

$$\frac{1}{2}[m^2(\psi_{xx}+\psi_{yy})+f]^2-\frac{m^4}{2}(\psi_{xx}-\psi_{yy})^2-2m^4\psi_{xy}^2+$$
$$m^2(\psi_x f_x+\psi_y f_y)-\left[m^2 g(z_{xx}+z_{yy})+\frac{1}{2}f^2\right]=0 \quad (3)$$

Equation (3) is a simple rearrangement of equation (2). For the purposes of the following discussion, it will be convenient to retain the balance equation in the special form (3). The boundary condition presently used is that suggested by Bolin [2]:

$$\frac{\partial\psi}{\partial s}=\frac{g}{f}\frac{\partial z}{\partial s}-\frac{\oint\frac{g}{f}dz}{\oint ds}$$

where s is distance measured counterclockwise along the boundary.

3. THE INITIAL PROGRAM FOR AUTOMATIC COMPUTING MACHINERY

The first method used on an operational basis was outlined in its essentials by the author [3]. It is a direct application of relaxation techniques [4] to the problem, and is similar to the method used by Bushby and Huckle [5] in their solutions of a modified balance equation.

One may write the finite-difference transformation of equation (3) in a rectangular mesh,

$$\eta^2-D_1^2-D_2^2+L-Z=0 \quad (4)$$

where η is absolute vorticity scaled so that the coefficient of the central value of ψ is -1, and D_1^2, D_2^2, L, and Z are appropriately scaled forms of the other terms in equation (3) in their order of appearance. It is to be noted that ψ at the central point is implicit only in the first term. It can be shown by variational calculus (e. g., Shuman [3]) that the differential equation (1) is elliptic if

$$g(z_{xx}+z_{yy})+\frac{1}{2}f^2-\psi_x f_x-\psi_y f_y>0$$

The corresponding condition for the finite difference equation (4) is probably

$$D_1^2+D_2^2-L+Z>0 \quad (5)$$

since no difficulty in treating it as a boundary-value problem has been encountered so long as the condition (5) is satisfied. However, since for meteorological data generally,

$$|Z|>>|L|$$

and since Z is almost exclusively positive, while D_1^2 and D_2^2 are positive definite, we first impose on the Z-field the condition

$$Z\geq 0 \quad (6)$$

by a technique that does not change its mean value nor, in practice, its large features significantly. The field of Z is scanned with a test for negative values. When a negative value of Z is encountered the values at the surrounding nearest four points are reduced by ¼ of the magnitude of Z at the central point, and the value of Z at the central point is increased to zero. Boundary values are excepted from change. A few scans are required to complete this operation.

In a numerical experiment involving some dozen cases, Mr. L. P. Carstensen of the development staff of our Unit has found that imposing the condition (6) implies only trivial changes in 500-mb. heights over areas of good data coverage, although changes of as much as 50 feet may be implied over areas of sparse data. It appears that the latter are due to analysis errors, due in turn to insufficient data.

Equation (4) may be written

$$\eta^\nu-[(D_1^\nu)^2+(D_2^\nu)^2-L^\nu+Z]^{1/2}=R^\nu \quad (7)$$

where ν is the scan count in the relaxation process, and R is a measure of the error of the current approximation. The Liebmann-type iteration is used, so that quantities designated by the superscript ν have only a brief existence. Specifically, the superscript ν denotes a value at a point during the ν-th scan after computation has been completed at the preceding point, but before computation has begun at the point in question. We will similarly write

$$\eta^{\nu+1}-[(D_1^\nu)^2+(D_2^\nu)^2-L^\nu+Z]^{1/2}=-\lambda R^\nu \quad (8)$$

Again, the quantity $\eta^{\nu+1}$ has only a brief existence. It is the value of η during the ν-th scan after computation has been completed at the point in question, but before computation at the succeeding point has begun. Thus,

$$\psi^{\nu+1}-\psi^\nu=-(\eta^{\nu+1}-\eta^\nu) \quad (9)$$

Equations (7), (8), and (9) may be combined.

$$\psi^{\nu+1}-\psi^\nu=(1+\lambda)\{\eta^\nu-[(D_1^\nu)^2+(D_2^\nu)^2-L^\nu+Z]^{1/2}\} \quad (10)$$

The quantity λ is the overrelaxation factor. It is to be noted that one cannot overrelax a residual based on

equation (4) as it stands, for with the little control one has over such an implied residual, and without D_1^2, D_2^2, L, or Z being bounded away from zero, one will surely encounter during the process imaginary values of η. On the other hand, the method of overrelaxation as indicated by equation (10) leads to no difficulty.

It has been found that the condition (6) is not always sufficient to satisfy the "elliptic" condition (5). At points where the other terms of the expression (5) are small, sometimes L is sufficiently positive to violate the condition. This has been handled by substituting zero for the expression (5) where the condition is violated, and keeping a count of such points during the scan. When convergence is otherwise indicated in four successive scans during which the condition (5) is violated at some points, the program proceeds in a similar fashion as before, but if, where negative, the absolute value of the expression (5) is smaller than a small number (an increment of Z is used, corresponding to $0.1\bar{f}^2$) it is ignored. The program has never failed to converge under these circumstances. The number of points involved has always been at least two orders of magnitude less than the total number (1,020) of points in the grid.

A disadvantage, from the viewpoint of time required, of the system (10) over solutions of linear equations is obvious, for it indicates a square root must be taken. The Newton iteration is used in our machine program. In terms of scan count, however, the system converges as fast as linear systems, as can be shown by solution of equation (3) treated as a Poisson equation in z. The total scan count is about the same whether one solves the equation for z with $\bar{f}g^{-1}\psi$ as the first guess, or whether one solves the equation for ψ with $\bar{f}^{-1}gz$ as the first guess. Total scan counts with such a first guess for our 30 x 34 grid of 381-km. mesh length is 100 to 125 scans. Our convergence criterion is

$$\bar{f}\,g^{-1}(\psi^{\nu+1}-\psi^{\nu}) \leq \tfrac{1}{4} \text{ foot}$$

where $\bar{f}$ is the Coriolis parameter at 45°. The program described above runs from 40 to 50 minutes on the Unit's IBM 701, including all input-output operations and other overhead.

4. A FAST METHOD

The method described in this section is, except in certain minor details, identical to a method arrived at independently by Miyakoda [6]. At the first writing of this paper, the author was unaware of Miyakoda's work.

The foregoing program was used on a daily operational basis from April 20, 1956 to January 30, 1957. On January 31, 1957 we instituted an operational code which cut the total running time to 25 minutes. More recently, we began using 12-hr. barotropic predictions of ψ as a first guess, which reduces running time to 18 minutes. In these later runs the relaxation itself is done in slightly more than 10 minutes, the rest of the time being occupied with input-output operations and other overhead, including checking and recovery procedures. In the new fast program, the square root indicated in equation (10) is taken only every 10 scans or so, it being otherwise held constant. We call this process "cyclic", a "cycle" consisting of the computations between the computation of the square root.[2] This procedure reduces the fundamental relaxation computation to the fastest possible; i. e., to that of a Poisson equation. The process may be stated symbolically,

$$(\psi^{\nu+1}-\psi^{\nu})^{\sigma+1}=(1+\lambda)\{\eta^{\nu,\sigma+1}-[(D_1^{\sigma})^2+(D_2^{\sigma})^2-L^{\sigma}+Z]^{1/2}\} \tag{11}$$

where σ is the cycle count. Instead of fixing the number of scans per cycle to a constant figure, which we now think would be a better procedure, we vary the convergence criterion, beginning with 64 ft. for the first cycle, and reduce it by one-half for each succeeding cycle until it reaches the value, ¼ ft., and then hold it constant. A cycle then consists of relaxation of a Poisson equation extending to indicated convergence. Ten to fifteen cycles are normally required for convergence, which is indicated by four successive cycles of one scan each. At the end, we have been requiring two scans indicating convergence with the older slower program described previously.

There have been occasional cases in which the new program has failed to converge. It had reached a near-perfect oscillation in which

$$\eta^{\sigma+1}-[(D_1^{\sigma})^2+(D_2^{\sigma})^2-L^{\sigma}+Z]^{1/2}=0$$

$$\eta^{\sigma+2}-[(D_1^{\sigma+1})^2+(D_2^{\sigma+1})^2-L^{\sigma+1}-Z]^{1/2}=0$$

This has been handled by halting the cyclic process after 32 cycles, and proceeding to convergence with the older program, which at that stage has taken less than 15 scans to accomplish. We have never encountered a case in which the new fast program diverges from the solution. The oscillation has been associated with small-scale irregularities in the analyses of z. When the field of z is treated to remove components with wavelengths of less than five grid increments, the oscillation does not occur. This is an empirically indicated way of avoiding oscillations in the process (11). It is thought that a better way would be to begin each cycle with a first guess which is a weighted average of the "solutions" from the previous two cycles. In this average one should weight the "solution" from the immediately preceding cycle very heavily, otherwise the rate of convergence would obviously be affected adversely.

5. REMARKS ON OTHER METHODS

One can linearize the balance equation before applying relaxation techniques, for example, as Bolin [2] has done.

[2] The cyclic concept is due to Mr. L. P. Carstensen, who first applied it to the balance equation in the form (1). He held constant during each cycle all terms except the first in equation (1). Some solutions were successfully obtained in this manner, but in some cases divergence from the solution was manifest. G. 'Arnason later showed on a theoretical basis the conditions under which such application of the cyclic procedure would diverge from the solution. Bushby and Huckle [5] reported failure for a similar scheme. 'Arnason's work has been submitted for publication.

His method is precisely equivalent to performing a single Newton iteration on the radical in equation (10) at each point during each scan. He sets $\lambda=0$. He has successfully solved the balance equation on a 500-point grid in this manner. As he has indicated his method as presented will fail for larger grids with strict convergence criteria unless he can arrive at a much better first guess of ψ than $\bar{f}^{-1}\, g\, z$. The reason for failure will be that as the grid becomes larger, the heights become a poorer estimate of the $\bar{f}\, g^{-1}\, \psi$-field, and overrelaxation will become necessary to keep running time down to acceptable limits. We have found that on a 19 x 29-point grid with a mesh length of 304.8 km., the balance equation can be solved in a reasonable time without overrelaxation, but on our operational 30 x 34-point grid with mesh length 381 km. it is not feasible to run the code to convergence from a first guess of $\bar{f}^{-1}\, g\, z$ without overrelaxation.

Due to the monotonic approach of the Newton iteration to the square root, Bolin's system, with $\lambda=0$, consistently under-relaxes. One could introduce overrelaxation into his system, but again due to the monotonic approach of the Newton iteration to the square root, such overrelaxation would not be overrelaxation in the usual sense. It is conceivable that one could correct this deficiency by adjusting the overrelaxation factor by empirical or statistical techniques, but here one is limited by the fact that if λ were to approach unity in his scheme, the Newton iteration itself would become divergent. An important test in this respect, is upon a trivial finite-difference field of one internal point. It is inconceivable that a system failing this simple test would succeed generally when applied to large fields.[3]

As mentioned previously, Bushby and Huckle [5] have successfully solved a modified balance equation, using a method essentially like that of our older program. Their grid was quite small, 256 points. It is in order to remark that they defined the balanced wind components as divergent, thus,

$$u=-f^{-1}\psi_y$$

$$v=+f^{-1}\psi_x$$

and derived a balance equation modified by the neglect of terms which inevitably arose from the space variation of f. The author has since pointed out elsewhere (Shuman [1]) that conventional use of the geostrophic approximation implies serious physical inconsistencies due to geostrophic divergence. The "balanced" wind fields of Bushby and Huckle largely retain the geostrophic divergence, so will produce no significant improvement in predictions over geostrophic wind fields.

In order to avoid explicitly the square root computation, Kasahara [7] has proposed the following cyclic procedure. He first sets down

$$(\nabla^2\psi^{\sigma+1}+f)^2-(D_1{}^\sigma)^2-(D_2{}^\sigma)^2+L^\sigma-Z=0$$

where σ is the cycle count. He then linearizes this equation by expanding the first term, thus,

$$[\nabla^2(\psi^{\sigma+1}-\psi^\sigma)]^2+2\,(\nabla^2\psi^\sigma+f)\nabla^2\,(\psi^{\sigma+1}-\psi^\sigma)+(\nabla^2\psi^\sigma+f)^2-$$

$$(D_1{}^\sigma)^2-(D_2{}^\sigma)^2+L^\sigma-Z=0$$

and neglecting the first term in the expansion. Fixing those terms superscribed by σ, he solves the resulting Poisson equation in $\psi^{\sigma+1}-\psi^\sigma$. He proceeds thus for several cycles to convergence. It is not clear that the division necessary to form the forcing functions for the Poisson equations will remain bounded, but if this were the case, and if the method were to prove otherwise universally convergent for meteorological data, it would indeed be of considerable interest.

ACKNOWLEDGMENTS

The author expresses his gratitude for the excellent services performed by the following former or present members of our staff of mathematicians: Lt. Cmdr. Albert Stickles, Mr. Stanley Herman, Mrs. Dorothy Hoover, Mrs. Miriam Bernhardt, and Mr. Otha Fuller.

REFERENCES

1. F. G. Shuman, "Predictive Consequences of Certain Physical Inconsistencies in the Geostrophic Barotropic Model," *Monthly Weather Review*, vol. 85 No. 7, July 1957, pp. 229–234.
2. B. Bolin, "An Improved Barotropic Model and Some Aspects of Using the Balance Equation for Three-Dimensional Flow," *Tellus*, vol. 8, No. 1, Feb. 1956, pp. 61–75.
3. F. G. Shuman, "A Method for Solving the Balance Equation," *Technical Memorandum* No. 6, Joint Numerical Weather Prediction Unit, 1955, 12 pp.
4. R. V. Southwell, *Relaxation Methods in Theoretical Physics*, Oxford, Clarendon Press, 1946, 248 pp.
5. F. H. Bushby and V. M. Huckle, "The Use of a Stream Function in a Two-Parameter Model of the Atmosphere," *Quarterly Journal of the Royal Meteorological Society*, vol. 82, No. 354, Oct. 1956 pp. 409–418.
6. K. Miyakoda, "On a Method of Solving the Balance Equation," *Journal of the Meteorological Society of Japan*, vol. 34, No. 6, Dec. 1956 pp. 364–367. Reprinted in *Vortex* No. 2, June 1957.
7. A. Kasahara, "A Method for Solving the Balance Equation with the Relative Vorticity as a Carrying Parameter," *Technical Report* No. 3 to the U. S. Weather Bureau (Contract Cwb 9016), Department of Meteorology, University of Chicago, 1957, 12 pp.

[3] G. 'Arnason in a paper soon to be published, has developed a method by a linearization technique which does not suffer from an implied monotonic approach to the square root. His technique, with a slower scan than the fast method outlined in this article, requires fewer scans. From a first guess of a quality similar to a 12-hr. prediction, his method competes successfully with the fast method described in section 4. Perhaps more important in terms of lasting interest, he has proven convergence for his method.